Toward an Intelligent and Friendly Eldercare Robot

Qinghao Zhuang[1,3], Yang Gao[1,3], Guangda Chen[1,2] and Xiaoping Chen[1,2]

***Abstract*— Nowadays, many countries are facing aging and aged populations and a shortage of eldercare resources, eldercare robots have been considered to solve this problem. This paper aims to describe the ideal and common application scenarios of eldercare robot and design a frame for hardware and software based on the application scenario. The main functions of the robot we designed include navigation, computer vision, object detection and manipulation, human expression analysis, dialogue management and so on. Besides, we will do some detailed analysis combine with common application scenarios based on our robot hardware and software design schemes to demonstrate our design.**

I. INTRODUCTION

In recent years, due to the miniaturization of family size and the busy work of offsprings, the proportion of older people living alone is increasing. The function of the traditional family raising the elderly is also affected. Although home care workers can provide services to the elderly, the number of workers is lacking, and it is difficult to find a suitable service provider. These issues have inspired researchers to research and develop universal robots that can enter the home and serve the elderly. Eldercare robots with the ability to sense, analyze, and execute will play a vital role in the daily lives of older people. Older people express their needs by talking to robots, which can be the help of everyday life or the treatment of missing mental and emotions. At the same time, the robot can also actively monitor the status of the elderly [1]. When the older people are in danger, the robot can perceive it and quickly determine how the danger should be resolved and then perform the action.

Currently, robotics has become more mature. However, there is no universal assistant robot that can enter the home. First of all, the family is a place full of uncertainty, which challenges the design and implementation of the unharmed eldercare robot. For robots, safety is the most fundamental guarantee, especially when serving the elderly. Secondly, the design of service robots is basically in the laboratory stage, and there are many special tricks for the environment, which cannot adapt to changes in the environment. Especially in the family, space is narrow, and the layout is semi-structured.

Based on the particularity of the eldercare robot and the uncertainty of the family environment, we proposed a basic work scene for the eldercare robot, including how the robot integrates into the family and how to provide better services for the elderly [2]. Besides, we designed hardware and software solutions for eldercare robots based on the described work scenarios and technologies used. In order to illustrate the rationality of the program, we carried out an example analysis based on specific scenarios and technologies.

The content of this article is as follows. We will describe the working scene of the eldercare robot in Section II. Second, the hardware and software design of the auxiliary robot is shown in Section III. Then we will analyze some scenarios in Section IV. Finally, we conclude in Section V.

1 Robotics Research Center, University of Science and Technology of China (USTC), Hefei, Anhui 230027, P.R. China. {sa517547, sa517085, cgdsss}@mail.ustc.edu.cn, {xpchen}@ustc.edu.cn.

2 School of Computer Science and Technology, USTC, Hefei, Anhui 230027, P.R. China.

3 School of Software Engineering, USTC, Hefei, Anhui 230051, P.R. China.

II. SCENE DESCRIPTION

The robot we envision is a robot that can move automatically, avoid obstacles, and interact with the environment and people. The environment in which the robot works should match the functions it has, especially the service robot. Service robots need to continually perceive the environment and interact with the environment during work. The feedback gained in the robotic environment will also help the robot's next move. The environment here includes furniture at home, older people, relatives of the elderly, pets, and more. Therefore, we define the robot, the environment, and the interaction between the two as a scenario. In this section, we mainly describe the use of aged care service robots, which involves the familiarity of the robot at the beginning of the family, the work process in the family, and the self-learning process [3].

A. First time into the family

Imagine a new service robot that just bought to home. As a new member of this family, it needs to understand and be familiar with the basic situation of this family [4]. For example, the layout of the house, where the living room is, where the bedroom is, and where it rests (charging). If the elder has some privacy and does not want the robot to enter, then the robot needs to be able to understand and remember this requirement. So the robot is required to have the necessary navigation functions and the ability to construct a semantic map.

After being familiar with the work environment, the robot needs to listen to the daily work requirements and arrangements of the elder. Of course, natural language dialogue is the most convenient and appropriate way. There should be no elder who wants to face the configuration interface of such a sophisticated robot.

Ideally, the robot should be able to initialize its configuration based on the personality of the elderly. It requires a robot with emotional recognition and multiple personality configurations. Once completed, the robot can start working.

B. Daily work

The ability of robots to work is the most important. The performance of workability is: (1) answer or complete the requirements of the elderly; (2) supervise the physical health of the elderly. If the elderly have abnormal heart rate, falls, unbearable pain or other dangerous situations, the robot needs to alert the family; (3) accompany the elder to spend every day with ease and pleasure, instead of waiting for the order and then making the cold respond. So the services provided by robots that we divided it into two parts: active services and passive services.

The active service must always be running while the robot is working. Active services mainly include monitoring the family environment and the situation of the elderly. When the furniture position changes, the robot should be able to discover and understand what is happening. When the elder does something, the robot should ask the elderly if they need help. At the same time, the robot itself should be able to manage itself, such as power monitoring, troubleshooting, online upgrades, and more.

Passive service is the interaction of the elderly or other people with the robot. The robot passively receives information and makes feedback or performs related actions. Passive services have a higher priority than active services. The robot should give immediate feedback when it receives a passive service request.

C. Learning ability

The behavior of the robot is not necessarily wholly correct. When the instructions are ambiguous, or the knowledge required to complete the task is insufficient, the robot may make a mistake or not know how to do it. Therefore, basic learning ability is an essential skill for service robots, and it is also a relatively common scene [3].

III. Hardware And Software Design

In this part, we will introduce the hardware design and software design and analyze some of the technical points.

First of all, Fig. 1 is the hardware and software logic architecture diagram of the robot we envisioned. From top to bottom, they are the skill layer, the decision layer, the function layer, the control layer, and the hardware layer. We introduce the order from hardware to software. The bottom layer is the hardware layer. This layer is the design of eldercare robot hardware. It contains the most basic hardware. Of course, if the robot needs to complete more complex functions, then additional interfaces should be kept for additional sensors. Above the hardware layer is the control layer, which is mainly the control of the robot hardware, including some hardware control like emergency stop and some software control like kinematics control.

Upward is the functional layer and the decision layer. These two layers are the main parts of our software program. The functional layer should contain various functions that the

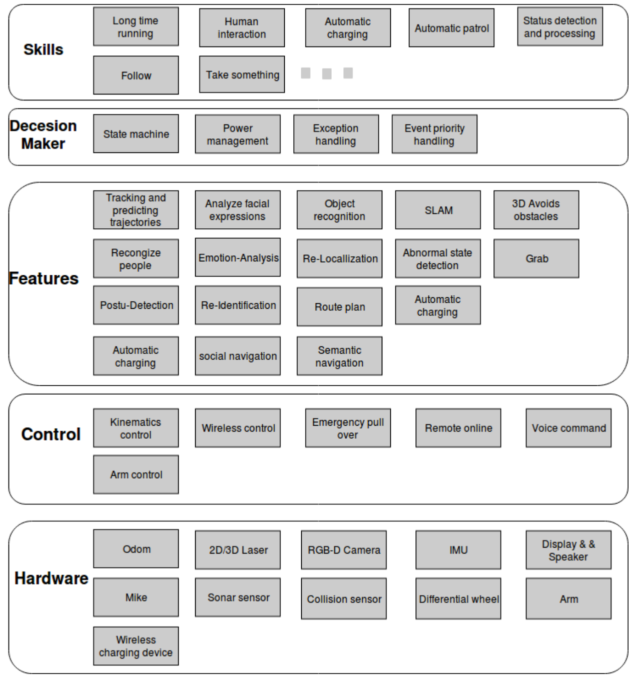

Fig. 1: Logical architecture of the eldercare robot.

robot's hardware can provide for the upper interface to call. The decision-making layer plays a crucial role in the order in which the robot needs to perform tasks in the environment and encounters the handling of emergencies or special events.

The uppermost skill layer is the task that the robot can perform in a specific scenario, mainly calling the lower interface to complete the various work, We will give detailed examples in the fourth part of the following, and we will not go into details here.

We will briefly summarize the technical points on the eldercare robot. Here we mainly introduce the following four technical points.

A. Self-localization and navigation

For self-localization and navigation, a 2D occupancy grid map is generated first from the raw data collected by laser scanners through a round journey within the rooms aforehand [5]. Then the map is manually annotated with the approximate location and area of rooms, doors, furniture and other interesting objects. Finally, a topological map will be automatically generated. 3D environment is also needed. The system receives the point cloud information from the camera device and then process the data with the localization provided by 2D grid map. Besides, sonar sensors can be used to detect glassware, which is useful for robots to avoid obstacles [6]. And multi-sensor spatial coordinate system conversion is well calibrated in advance [7].

B. Computer vision

In our system, two cameras are needed, a high-resolution CCD camera and an RGB-D camera, to obtain aligned RGB-D images as well as high-quality RGB images. Both cameras need to calibrate so we can directly get the correspondence between the images. We obtain an aligned RGB-D image by combining the RGB image with the depth image. With such an aligned RGB-D image, our vision module is capable of detecting, tracking people, recognizing different kinds of objects and even to recognizing people's facial expression [8].

C. Object grasping

For eldercare robot, object grasping function is a necessary one. In this part, we decompose the grasping problem into multiple sequential task state: approach, establish grasp, move object, release and retract. Each governed by either local control, continuous optimization, or some combination thereof [9]. In the grasping process, the robot's arm is controlled by the Decision Maker layer and executed by the Control layer [10].

D. Human expression analysis

The robot needs to observe the human's expression and analyze the observed facial expressions of the person. For this purpose, MIT's method [11] is a good choice. In this part, Our main task is mainly to extract facial features and analyze the extracted features to obtain a preliminary human emotion estimate. This emotion we got here is not the final answer, and we will combine the dialogue analysis to get a better emotion estimate of the human. Our ultimate goal is to discover the real-time status of the elderly and take appropriate measures [12].

E. Dialogue management

The robots Dialogue Understanding module for Human-Robot Interaction contains Speech Recognition module and Natural Language Understanding module, and it provides the interface for communication between users and the robot. This part should be able to different synthesis languages to different people. As for recognition, a configuration represented by BNF grammar is required. Since each test has its own set of possible speech commands, we pre-build several configurations to include all the possible commands for each test. The Natural Language Understanding module is used for the translation to its semantic representation. With the results of Speech Recognition module and the semantic information of the speech, the Natural Language Understanding module can update the World Model, which contains the information from the perceptual model of the robots internal state, and/or invoke the Task Planning [13] module for fulfilling the task. The translation from the results of the Speech Recognition module to semantic representation consists of the syntactic parsing and the semantic interpretation. For the syntactic parsing, we recommend using the Stanford parser [14] to obtain the syntax tree of the speech. For the semantic interpretation, the lambda-calculus [15] is applied to the syntax tree to construct the semantics. Fig. 2 shows an example of a semantic interpretation.

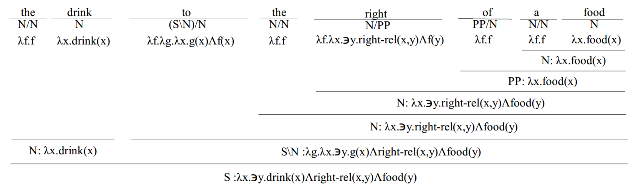

Fig. 2: Example parse of the drink to the right of a food. The first row of the derivation retrieves lexical categories from the lexicon, while the remaining rows represent applications of CCG combinators.

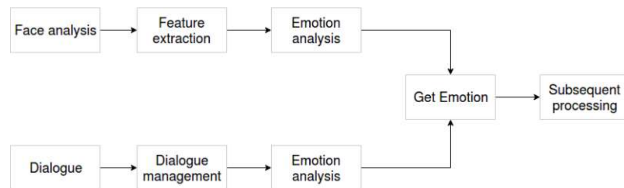

Fig. 3: Flow chart of multimodal sentiment analysis. A variety of information sources to analyze the current state of the elderly, and weight each estimate to more accurately obtain the current state of the elderly

F. Multimodal sentiment analysis

Here we combine the aforementioned dialogue management and human expression analysis to combine the two results in a weighted manner to determine the current emotional state of the person and take corresponding measures which is shown in Fig. 3.

G. Task planning and task priority order setting

The robot should have the task of planning and prioritizing tasks. We add a priority classifier for task planning based on a good mission plan to determine the priority of things that should be handled in the environment. For example, the old man issued an order to "give me a glass of water". After that, the old man fell down and robot should immediately suspend the action of taking water, and then deal with the higher priority incident of "the old man fell down". The process is shown in Fig. 4.

IV. Scene Analysis

In this part, we will analyze and describe the various functions of the robot in three ways: active service, passive service, and state management. We will give some examples and describe what kind of actions the robot should do or how to interact with people in these examples.

A. Active service

The ability of many older people to remember is getting worse. When the elderly are alone at home, they often forget to turn off the equipment, such as gas, faucets, appliances, and more. There is a great potential danger. We require robots to proactively analyze the environment within the home during operation to monitor for potential hazards. Considering that the definition of danger is relative in different environments, for which conditions are dangerous

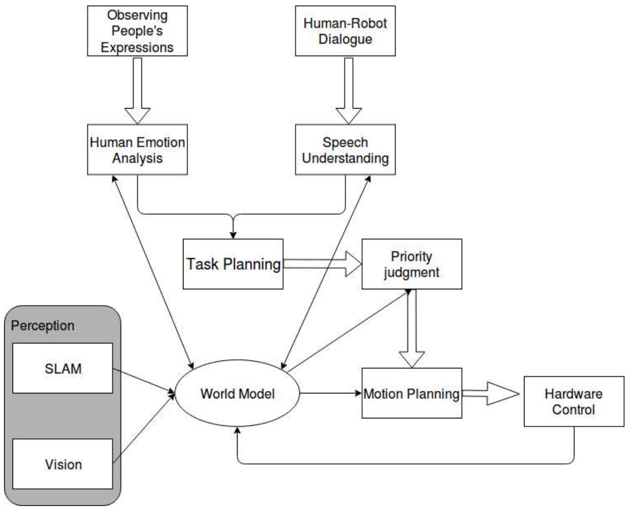

Fig. 4: Flow chart of task planning and task priority order setting. The current state of the elderly is analyzed and the mission planning is performed by the detected voice and image of the elderly, and the task priority is determined during the mission planning process.

and which are not dangerous, the robot needs to learn, that is, to be informed.

The sudden situation of the elderly can be used to remind the family or seek professional help. For example, the elder suddenly falls into the home, or the robot can quickly identify and handle other situations. If the elder falls, he can quickly find and call the doctor for help.

Through interaction with the elderly, the robot can feel the state of the elderly and adjust the personality configuration. When the elderly feel depressed or other mood changes, the robot can identify and process the corresponding treatment.

B. Passive service

The task that the robot completes under the exact command of the person.For example, the elderly will ask the robot to help pick things up or other things. The robot can clearly recognize the human's command and react quickly, while the task can't be completed by itself, the robot could ask for help or feedback to the serviced person.

C. State management and priority set

The robot has the ability to self-test and self-repair when an abnormal state occurs and can handle it when there are some simple hardware or software problems. When a big problem occurs, the robot can remotely generate a report and transmit it to a professional maintenance person.

The robot manages its power and can automatically charge, and this charging time is scheduled to be the time of rest of the serviced person.

The robot should set a severe priority of the various state in the environment. For example, when the old man tells the robot, "give me a glass of water," after that, the old man falls. The robot can judge "old man falls" is a higher priority than "give me a glass of water."

V. CONCLUSIONS

In the above, we made some conjectures and designs about the eldercare robot. Some of these technologies have been implemented and applied. Some technologies have cited new research results, and may not be practical. In any case, the actual application of the eldercare robot needs some development time. We hope that this future pension robot can make significant progress and help those who need it.

REFERENCES

[1] K. Otsu, H. Fukuda, Y. Kobayashi, and Y. Kuno, "Remote monitoring and communication system with a doll-like robot for the elderly," in *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2016, pp. 5941–5946.

[2] D. Portugal, P. Trindade, E. Christodoulou, G. Samaras, and J. Dias, "On the development of a service robot for social interaction with the elderly," 2015.

[3] W. YINGZI and Z. MINGYANG, "A framework for the robot skill learning using reinforcement learning," in *SPIE proceedings series*. Society of Photo-Optical Instrumentation Engineers, 2003, pp. 910–914.

[4] W. Shuai and X.-p. Chen, "Kejia: towards an autonomous service robot with tolerance of unexpected environmental changes," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 3, pp. 307–317, 2019.

[5] G. Grisetti, C. Stachniss, W. Burgard *et al.*, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE transactions on Robotics*, vol. 23, no. 1, p. 34, 2007.

[6] K. Okuda, M. Miyake, H. Takai, and K. Tachibana, "Obstacle arrangement detection using multichannel ultrasonic sonar for indoor mobile robots," *Artificial Life and Robotics*, vol. 15, no. 2, pp. 229–233, 2010.

[7] G. Chen, G. Cui, Z. Jin, F. Wu, and X. Chen, "Accurate intrinsic and extrinsic calibration of RGB-d cameras with GP-based depth correction," *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2685–2694, apr 2019. [Online]. Available: https://doi.org/10.1109%2Fjsen.2018.2889805

[8] J. Liu, Z. Zhang, B. Tang, and X. Chen, "Wrighteagle@ home 2017 team description paper," *RoboCup@ Home 2017 Team Description Papers*, 2017.

[9] D. Kappler, F. Meier, J. Issac, J. Mainprice, C. G. Cifuentes, M. Wüthrich, V. Berenz, S. Schaal, N. Ratliff, and J. Bohg, "Real-time perception meets reactive motion generation," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1864–1871, 2018.

[10] Y. Tao, H. Liu, X. Deng, Y. Chen, H. Xiong, Z. Fang, X. Xie, and X. Xu, "A robot self-learning grasping control method based on gaussian process and bayesian algorithm," in *Transactions on Intelligent Welding Manufacturing*. Springer, 2018, pp. 79–98.

[11] M. Feffer, R. W. Picard *et al.*, "A mixture of personalized experts for human affect estimation," in *International Conference on Machine Learning and Data Mining in Pattern Recognition*. Springer, 2018, pp. 316–330.

[12] S. A. Taylor, N. Jaques, E. Nosakhare, A. Sano, and R. Picard, "Personalized multitask learning for predicting tomorrow's mood, stress, and health," *IEEE Transactions on Affective Computing*, 2017.

[13] V. Lifschitz, "Answer set planning," in *International Conference on Logic Programming and Nonmonotonic Reasoning*. Springer, 1999, pp. 373–374.

[14] D. Klein and C. D. Manning, "Accurate unlexicalized parsing," in *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*. Association for Computational Linguistics, 2003, pp. 423–430.

[15] P. Blackburn, "Representation and inference for natural language: A first course in computational semantics," 2005.